

Measuring Spike Train Correlation with Non-Parametric Statistics Coefficient

J. H Soletta, F. D. Farfán and C. J. Felice

Abstract— Measure correlation between spike trains is a fundamental step for the study of neural systems. There are many alternatives to measure correlation, but not all possess the required properties. In this paper we propose to use non-parametric coefficients of correlation, coefficients Spearman and Kendall. To analyze their properties were generated computationally trains of spikes that simulate different experimental conditions, then the proposed coefficients were calculated and compared with the Pearson coefficient. The results show that under certain experimental conditions Kendall coefficient is more appropriate to quantify correlations between spikes trains.

Keywords— Neural Correlation, Spike Train, Kendall Coefficient, Spearman Coefficient.

I. INTRODUCCION

LA COMUNICACIÓN entre neuronas está basada en potenciales de acción que ocurren en el transcurso del tiempo a lo cual llamamos trenes de espigas. La sincronía o correlación entre trenes de espigas de distintas neuronas está asociada a diversos procesos biológicos como: codificación de estímulos, atención, discriminación de estímulos [1], desarrollo de interconexiones neuronales [2], etc. Cuantificar de manera apropiada estas correlaciones es crucial para el entendimiento de muchos procesos biológicos asociados a la comunicación entre neuronas.

Actualmente existen diversas alternativas para cuantificar correlaciones entre dos trenes de espigas [3] [4] [5]. Una medida ampliamente utilizada es el coeficiente de correlación de Pearson aplicado al conteo de espigas, pero se ha demostrado que su valor se modifica con la frecuencia de disparo de espigas [1] lo cual es una característica no deseada. Adicionalmente el coeficiente de Pearson solo mide el grado de asociación lineal entre las señales consideradas, por lo tanto si las señales que se analizan se asocian de manera no lineal no es recomendable utilizar este coeficiente.

Una alternativa estadística para medir correlación es utilizar coeficientes de correlación no- paramétricos. Estos coeficientes poseen la propiedad de poder captar asociaciones

distintas a la lineal y de no necesitar una gran cantidad de mediciones para obtener un valor confiable de correlación [6]. En este trabajo proponemos utilizar dos métodos para medir correlación basada en estadística no- paramétrica, el coeficiente de correlación de Kendall y el coeficiente de correlación de Spearman. Para analizar la eficiencia de estos métodos se generaron trenes de espigas de manera artificial y se utilizaron los criterios propuestos por [7]. Los resultados muestran que bajo condiciones experimentales específicas el coeficiente de Kendall mide de manera más apropiada la correlación entre trenes de espigas que el coeficiente de Pearson.

II. MATERIALES Y MÉTODOS

Un requisito previo para aplicar las mediciones de correlación que se analizarán es calcular el conteo de espigas y posteriormente ordenar los vectores resultantes por rangos.

Conteo de espigas: A los trenes de espigas registrados a lo largo de un tiempo T , se los divide en $N=T/d$ bins de ancho d . Luego se procede a contar la cantidad de espigas presentes dentro de cada bin, obteniendo los vectores X e Y de longitud N . El elemento X_i representa la cantidad de espigas presentes en el bin i .

Rangos de un vector: consiste en asignar un número a cada elemento de un vector según el valor que poseen, de esta manera al elemento de menor valor se le asigna el rango 1, rango 2 al elemento siguiente, etc.

Medición de Correlación

Coeficiente de Pearson: La correlación entre los trenes de espigas X e Y viene dado por:

$$C_{xy} = \frac{\frac{1}{N} \sum_{i=1}^N (X_i - X_m)(Y_i - Y_m)}{\sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - X_m)^2} \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - Y_m)^2}} \quad (1)$$

Donde X_m e Y_m son los valores medios de X e Y respectivamente, X_i es el número de espigas en el bin i .

Correlación de Kendall: Para calcular el coeficiente de correlación de Kendall se debe reemplazar los valores de cada bin por su respectivo rango, posteriormente se reordenan los elementos de los vectores X e Y siguiendo el orden de los rangos de unos de ellos. Luego se procede a analizar el número de concordancias en el nuevo ordenamiento de los vectores. El Coeficiente de correlación de Kendall viene dado por:

$$T = \frac{\sum_{i=1}^N \sum_{j=1}^N \text{sgn}(X_i - X_j) \text{sgn}(Y_i - Y_j)}{N(N-1)} \quad (2)$$

donde

J. H. Soletta, Laboratorio de Medios e Interfases (LAMEIN), Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, jorge.soletta@gmail.com

F. D. Farfán, Laboratorio de Medios e Interfases (LAMEIN), Facultad de Ciencias Exactas y Tecnología Universidad Nacional de Tucumán, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, ffarfan@herrera.unt.edu.ar

C. J. Felice, Laboratorio de Medios e Interfases (LAMEIN), Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, cfelice@herrera.unt.edu.ar

$$\text{sgn}(Z_i - Z_j) = \begin{cases} 1 & \text{si } (Z_i - Z_j) > 0 \\ 0 & \text{si } (Z_i - Z_j) = 0 \\ -1 & \text{si } (Z_i - Z_j) < 0 \end{cases} \quad (3)$$

Donde N es el número de bins.

Correlación de Spearman: El coeficiente de correlación viene dado por

$$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N^3 - N} \quad (4)$$

Donde d_i es la diferencia entre los rangos asignados a X_i e Y_i y N es la cantidad de bins de los trenes de espigas.

Generación de trenes de espigas

Para evaluar las mediciones de correlación se utilizaron dos modelos para generar trenes de espigas mediante simulación. El primer modelo que se implementó es el de Poisson y los valores de sus parámetros fueron:

- Frecuencia de disparo (λ): 0.1-5 Hz
- Tiempo de registro (T): 50- 800 s
- Ventana de sincronía 0.001 s

El segundo modelo utilizado es el propuesto por Macke [8], el cual permite generar trenes de espigas con un determinado valor de correlación.

Evaluación de las mediciones

Las mediciones fueron analizadas utilizando los criterios propuestos por Cutts y Eglén [7]. De manera breve se enuncian los criterios de evaluación utilizados.

- Robustez ante la variación en la frecuencia de disparo. La prueba consiste en calcular la autocorrelación de un tren de espigas haciendo variar la frecuencia de disparo y manteniendo constantes los otros parámetros. Se espera que la autocorrelación sea independiente de la frecuencia de disparo.
- Dependencia de la frecuencia. En esta prueba se calcula la correlación entre dos trenes de espigas, uno con frecuencia de disparo fija (λ_1) y el otro con frecuencia de disparo variable (λ_2). Se esperaría que si $\lambda_1 \equiv \lambda_2$ la correlación posea un valor constante, y que este valor disminuya medida que λ_2 difiera de λ_1 .
- Robustez a la cantidad de datos. Esta prueba consiste en calcular la correlación entre dos trenes de espigas para distintos valores de longitud (T) del tren de espigas, en este caso ambos trenes de espigas son generados utilizando la misma frecuencia de disparo λ .
- Robustez ante variaciones del ancho de bin (d). Se generaron dos trenes de espigas con la misma frecuencia de disparo y se calculó la correlación entre ellos utilizando distintos valores de d . Pequeñas variaciones en d no deberían provocar cambios significativos en la medición.
- Discriminación entre correlación y no-correlación. Utilizando el modelo 2 se generaron dos trenes de espigas que no poseen correlación entre ellos. Se espera que las mediciones de correlación sean próximas a cero.

- Robustez ante la presencia de periodos inactivos. Los periodos en los cuales ambos trenes de espigas están inactivos (no presentan espigas) no deberían modificar el valor de la medición. En este ensayo ambos trenes de espigas fueron generados utilizando dos frecuencias, en la primera mitad del registro se utilizó λ_1 y en la segunda mitad λ_2 . Para modelar los periodos sin espigas λ_2 tomó valores comprendidos entre $0.1 \leq \lambda_2 \leq \lambda_1$, de esta manera se modeló distintos grados de periodos inactivos (Fig. 1).

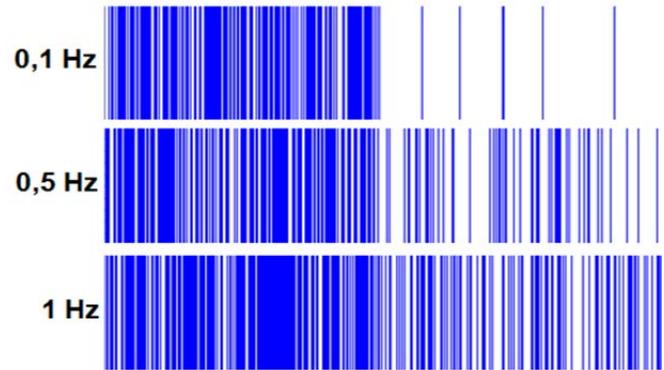


Figura 1. Trenes de espigas generados para el ensayo de robustez ante la presencia de periodos inactivos. La segunda mitad del tren de espiga posee diferentes valores de frecuencia de disparo, en los paneles superior, medio e inferior la frecuencia de disparo de la segunda mitad es 0,1; 0,5 y 1 Hz respectivamente.

III. RESULTADOS

Robustez ante la variación en la frecuencia de disparo (Autocorrelación). En la Fig. 2 se observa que los coeficientes de correlación de Pearson y Spearman permanecen invariantes e iguales a 1 cuando se calcula autocorrelación a diferentes frecuencias de disparo de espigas. A diferencia de los dos anteriores el coeficiente de Kendall muestra una variación en su valor en función de la frecuencia, para valores de frecuencia bajos (0,1 Hz) la medición es de 0,5 aproximadamente y tiende a aumentar a medida que aumenta la frecuencia de disparo.

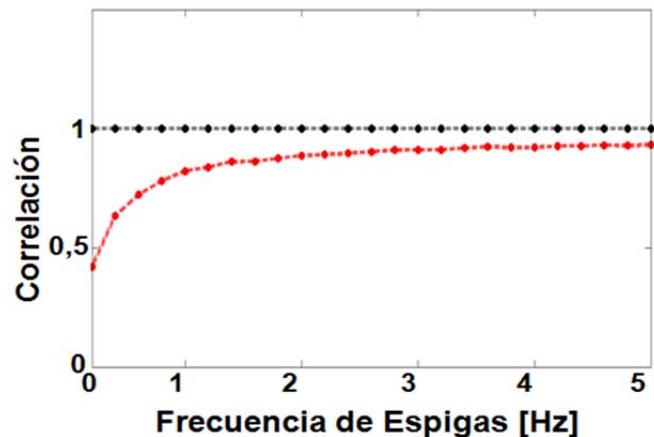


Figura 2. Autocorrelación de un tren de espigas calculado para diferentes frecuencias de disparos. Los puntos corresponden al valor promedio de 10 repeticiones de la medición. Las curvas de color azul, rojo y negro corresponden a los coeficientes de Pearson, Kendall y Spearman respectivamente.

Dependencia de la frecuencia (una fija y la otra variable).

En la Fig. 3 se observa que los valores de correlación para los tres coeficientes toman valores comprendidos entre -0,1 y 0,1 cuando la frecuencia de disparo del segundo tren (λ_2) varía entre 1 y 6 Hz. Este comportamiento se mantiene para los coeficientes de Pearson y Kendall en el rango de frecuencia de 0,1-1 Hz, pero a diferencia de estos dos el coeficiente de Spearman posee valores superiores en este rango de frecuencia a los que presenta en el resto del registro.

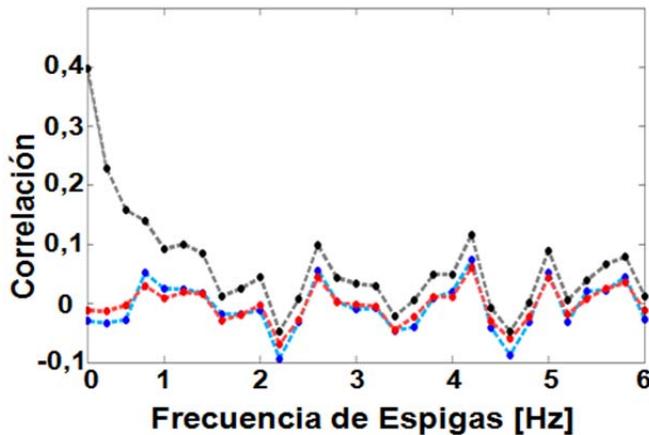


Figura 3. Mediciones de correlación entre dos trenes de espigas generados a dos frecuencias de disparos diferentes. El primer tren de espigas posee una frecuencia de disparo de 3 Hz y el segundo una frecuencia de disparo variable entre (0,1- 6Hz). Los puntos corresponden al valor promedio de 10 repeticiones de la medición. Las curvas de color azul, rojo y negro corresponden a los coeficientes de Pearson, Kendall y Spearman respectivamente.

Robustez a la cantidad de datos. Para analizar este ítem se calculó la correlación entre dos trenes de espigas con longitudes de registros comprendidos entre 50 y 800 segundos. En la Fig 4 se observa que las mediciones obtenidas con los tres coeficientes se comportan de manera similar. Para longitudes de registros comprendidos entre 50 y 400 segundos los valores de los coeficientes poseen una gran variación, la cual tiende a disminuir a medida que aumenta la longitud del registro.

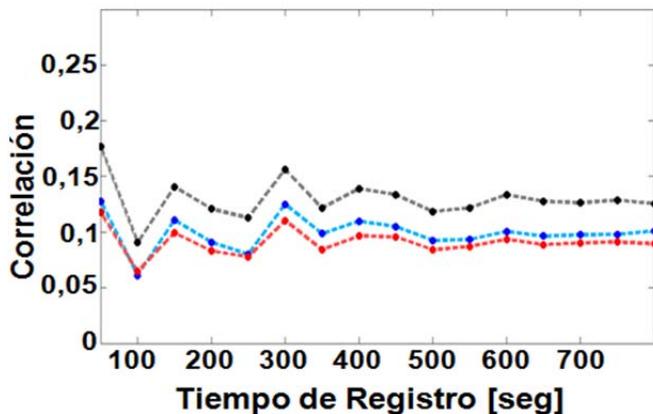


Figura 4. Mediciones de correlación entre dos trenes de espigas generados con la misma frecuencia de disparo (3 Hz). La longitud del registro varía entre 50-800 segundos. Los puntos corresponden al valor promedio de 10 repeticiones de la medición. Las curvas de color azul, rojo y negro corresponden a los coeficientes de Pearson, Kendall y Spearman respectivamente.

Robustez ante pequeñas variaciones en el ancho de bins (d).

El ancho de los bins afecta a las mediciones de correlación dependiendo de la frecuencia de disparo de los trenes de espigas. En la Fig 5- A se observa que el valor de d afecta al valor del coeficiente de Spearman, mientras que los otros coeficientes permanecen invariantes cuando la frecuencia de disparo es de 0,2 Hz. En la Fig. 5-B se observa que el valor de d no afecta a los valores de los coeficientes cuando la frecuencia de disparo es de 3 Hz.

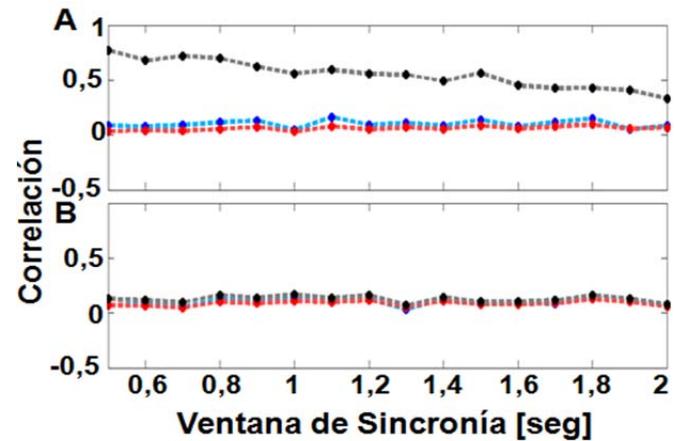


Figura 5. Mediciones de correlación entre dos trenes de espigas generados a una misma frecuencia de disparo. En cada caso se calculó la correlación haciendo variar el ancho de (d) los bins. Los puntos corresponden al valor promedio de 10 repeticiones de la medición. Las curvas de color azul, rojo y negro corresponden a los coeficientes de Pearson, Kendall y Spearman respectivamente. (A) Los trenes de espigas fueron generados con una frecuencia de disparo de 0,2 Hz. (B) Los trenes fueron generados utilizando una frecuencia de disparo de 3 Hz.

Medición de no- correlación. Cuando dos trenes de espigas son generados sin correlación las mediciones del coeficiente de Pearson toman en promedio valores más cercanos a cero, en comparación con los otros, ver Fig 6. Se debe notar que es el coeficiente de Kendall el que posee menor dispersión en la medición.

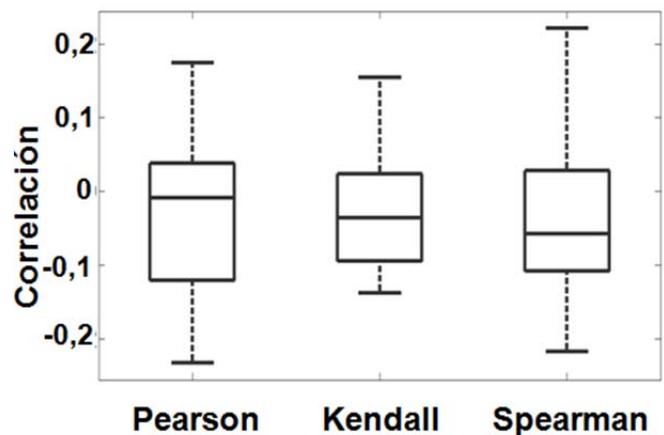


Figura 6. Diagramas de Caja de las mediciones de correlación entre dos trenes de espigas generados con valor de correlación igual a cero.

Robustez ante periodos inactivos. El coeficiente de Kendall es el que posee menor valor cuando se realiza el ensayo de robustez ante periodos inactivos, ver Fig 7. La diferencia entre

los valores de los coeficientes tiende a disminuir a medida que aumenta la frecuencia de disparo λ_2 .

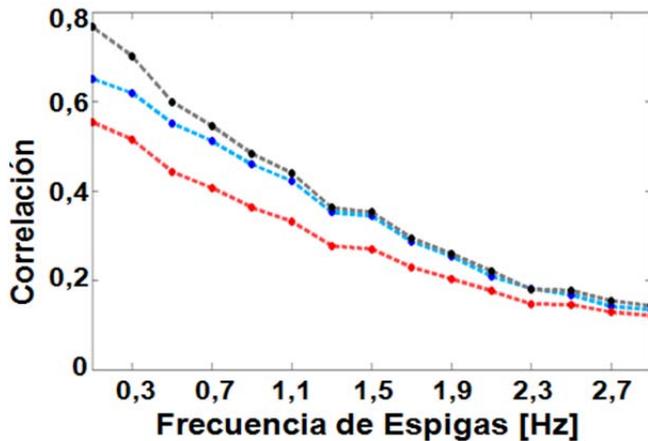


Figura 7. Mediciones de correlación entre dos trenes de espigas. La primera mitad de los trenes de espigas poseen una frecuencia de disparo de 3 Hz y la segunda mitad una frecuencia de disparo variable entre (0.1- 3Hz). Los puntos corresponden al valor promedio de 10 repeticiones de la medición. Las curvas de color azul, rojo y negro corresponden a los coeficientes de Pearson, Kendall y Spearman respectivamente.

IV. DISCUSIÓN Y CONCLUSIÓN

En este trabajo se propuso medir el grado de asociación entre dos trenes de espigas utilizando métodos estadísticos no-paramétricos. Los valores de correlación obtenidos con los coeficientes de Kendall y Spearman en general son similares a valores obtenidos con el coeficiente de correlación de Pearson, pero poseen algunas propiedades destacables.

La medición de correlación de los tres coeficientes se ve afectada por la longitud que poseen los trenes de espigas. Para obtener valores estables de correlación durante situaciones experimentales la longitud de los registros deberán poseer como mínimo 6,6 minutos (Fig. 5). Este valor mínimo de longitud de registro difiere del utilizado por algunos autores los cuales proponen un valor de longitud usual de 2 minutos [9].

Los valores de correlación obtenidos con el coeficiente de Kendall poseen menor dispersión que los valores obtenidos con coeficiente de Pearson y además es más robusto ante la presencia de periodos sin espigas (Fig 6 y 7). Estas características lo convierten en un método de medición útil cuando se desea analizar asociación entre neuronas que son sometidas a un mismo estímulo.

AGRADECIMIENTOS

Este trabajo ha sido financiado por fondos de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCYT); Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), y el Consejo de Investigaciones de la Universidad Nacional de Tucumán (CIUNT), también con fondos institucionales del Instituto Superior de Investigaciones Biológicas (INSIBIO).

REFERENCIAS

[1] De la Rocha J, et al. Correlation between neural spike trains increases with firing rate, *Nature* 448:802-806, 2007.

- [2] Stafford BK, Sher A, Litke AM, Feldheim DA, Spatial-temporal patterns of retinal waves underlying activity-dependent refinement of retinofugal projections. *Neuron* 64:200–212, 2009.
- [3] Wong RO, Meister M, Shatz CJ Transient period of correlated bursting activity during development of the mammalian retina. *Neuron* 11: 923–938, 1993.
- [4] Kruskal PB, Stanis JJ, McNaughton BL, Thomas PJ, A binless correlation measure reduces the variability of memory reactivation estimates. *Stat Med* 26:3997–4008, 2007.
- [5] Kerschensteiner D, Wong RO, A precisely timed asynchronous pattern of ON and OFF retinal ganglion cell activity during propagation of retinal waves. *Neuron* 58:851–858, 2008.
- [6] Walpole E R, Myers R H, Myers S L, Probabilidad y Estadística para Ingenieros, 6ta edición
- [7] Cutts C and Eglén S, Detecting Pairwise Correlations in Spike Trains: An Objective Comparison of Methods and Application to the Study of Retinal Waves, *The Journal of Neuroscience* 34(43): 14288-14303, 2014
- [8] Macke J, et al. Generating Spike Trains with Specified Correlation Coefficients, *Neural Computation* 21, 397–423, 2009.
- [9] Eglén SJ, Weeks M, Jessop M, Simonotto J, Jackson T, Sernagor E. A data repository and analysis framework for spontaneous neural activity recordings in developing retina. *GigaScience* 3:3, 2014



Jorge H. Soletta recibió el título de Ingeniero Biomédico en la Universidad Nacional de Tucumán (UNT) en el año 2009. Actualmente cursa sus estudios de Doctorado (PhD) en UNT. Sus áreas de investigación incluyen Procesamiento de Imágenes y Señales, Vision Humana y Neurociencias.



F. D. Farfán recibió los títulos de Ingeniero Electrónico, Magíster en Bioingeniería y Doctor en Ciencias Biológicas en la Universidad Nacional de Tucumán (UNT). Es docente en la UNT e Investigador en el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Sus áreas de investigación son las Neurociencias, Bioingeniería y Biomimética.



de tejidos.

C. J. Felice recibió los títulos de Ingeniero Electrónico y Doctor en Bioingeniería en la Universidad Nacional de Tucumán (UNT). Es director del Laboratorio de Medios e Interfaces, profesor en la UNT e Investigador Principal en el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Investiga en bioimpedancia y dirige un grupo de investigadores en biosensores, biomimética e ingeniería